JOURNAL OF VIROLOGY, Oct. 1999, p. 8349–8355 Vol. 73, No. 10
0022-538X/99/$04.00+0

A Phylogenetically Conserved Hairpin-Type 3′ Untranslated Region Pseudoknot Functions in Coronavirus RNA Replication

GWYN D. WILLIAMS, RUEY-YI CHANG,† AND DAVID A. BRIAN*

Department of Microbiology, University of Tennessee, Knoxville, Tennessee 37996-0845

Received 12 April 1999/Accepted 2 July 1999

Secondary and tertiary structures in the 3′ untranslated region (UTR) of plus-strand RNA viruses have been postulated to function as control elements in RNA replication, transcription, and translation. Here we describe a 54-nucleotide (nt) hairpin-type pseudoknot within the 288-nt 3′ UTR of the bovine coronavirus genome and show by mutational analysis of both stems that the pseudoknotted structure is required for the replication of a defective interfering RNA genome. The pseudoknot is phylogenetically conserved among coronaviruses both in location and in shape but only partially in nucleotide sequence, and evolutionary covariation of bases to maintain G · U pairings indicates that it functions in the plus strand. RNase probing of synthetic transcripts provided additional evidence of its tertiary structure and also identified the possible existence of two conformational states. These results indicate that the 3′ UTR pseudoknot is involved in coronavirus RNA replication and lead us to postulate that it functions as a regulatory control element.

Coronaviruses are enveloped, nonsegmented, intracytoplasmically replicating, plus-strand RNA viruses with the largest known viral RNA genome (approximately 30 kb) (reviewed in references 31 and 38). During replication, a 3′-coterminal nested set of subgenomic mRNAs are synthesized by an unresolved mechanism of discontinuous transcription (3, 51, 61) that places a common 5′-terminal leader sequence (only part of the 5′ untranslated region [UTR]) on each mRNA. It has been suggested that the common 5′ and 3′ termini on genomic and subgenomic mRNAs enable these molecules to amplify by a replication mechanism (57). Such a pathway would explain the existence of minus-strand copies of subgenomic mRNAs (20, 21, 56, 57) and of subgenomic mRNA-length replicative intermediates (50, 52). Replication of coronavirus RNA molecules from input plus strands, however, has been demonstrated for only the viral genome (when extracted from virions and transfected into uninfected cells [7, 53]) and defective interfering (DI) RNAs (when synthesized in vitro from cDNA clones and transfected into helper virus-infected cells [11, 29]), leaving unresolved by direct proof the degree of replicability of coronavirus subgenomic mRNAs. If coronavirus subgenomic mRNAs are lacking in signals for replication, it is unlikely that they would map within the 3′ UTR since this region is identical among the genome and subgenomic mRNAs (8).

Several reports have provided evidence for higher-order structural elements in the 3′ UTRs of plus-strand RNA viruses that are thought to function in RNA replication or translation by interacting with viral or cellular proteins. Stem-loop structures in the 3′ UTR of rubella virus (42), West Nile virus (4), and hepatitis C virus (5, 25) represent recognition sites for cellular proteins. The recruitment of a *trans*-acting factor(s) to postulated higher-order structures in other animal viruses including picornaviruses (26, 40, 41, 45, 46, 64), flaviviruses (59), and coronaviruses (24) has not been elucidated. Among plant viruses, alfalfa mosaic virus coat protein binds a 3′ UTR with extensive secondary structure (2) and cellular proteins recognize higher-order motifs within the 3′ termini (14, 15, 19) and in upstream sequences (33). The functional relevance of a tertiary interaction involving the poly(A) tail and 3′ UTR of bamboo mosaic virus has not yet been characterized (65).

The use of coronavirus DI RNAs to define replication signaling elements within genomic termini would seem reasonable since the 5′ and 3′ UTRs on known DI RNAs, the only available cloned coronavirus replicons, are identical to those on cognate virus genomes (8). Thus, in the bovine coronavirus (BCV) genome and DI RNA, the 5′ UTR is 210 nucleotides (nt) in length, inclusive of the common 65-nt leader sequence, and the 3′ UTR is 288 nt, exclusive of the poly(A) tail (11, 32). To investigate the potential for signaling structural elements within the BCV 3′ UTR, a computer-based RNA-folding algorithm was used to predict thermodynamically stable secondary structures. We then analyzed a potential pseudoknot beginning 63 nt downstream from the termination codon of the N gene by genetic and biochemical approaches, including phylogenetic comparisons, site-directed mutagenesis in a cloned DI RNA followed by Northern blotting, and enzymatic structure probing of in vitro-transcribed RNA. These findings strongly suggest that the pseudoknot plays a role in coronavirus RNA replication.

MATERIALS AND METHODS

Computer prediction of secondary structure. The Microgenie sequence analysis program (Beckman Instruments) was used to identify inverted repeats and to calculate the free energy of secondary structures by following the rules of Tinoco (63).

Virus and cells. A DI RNA-free stock of the Mebus strain of BCV at 4.5×10^8 PFU/ml was prepared and used as a helper virus (11). The human rectal tumor cell line HRT-18 was used in all experiments (11).

Construction of mutant DI RNAs. Mutant DI RNAs were all modifications of pDrep1, a cloned naturally occurring DI RNA modified to contain a 30-nt reporter insert (Fig. 1A) (11). Mutations in the two stems of the pseudoknot were introduced by a method based on the technique of gene splicing by overlap extension (23, 55). The restriction endonuclease sites used in making the mutant DI RNAs are shown in Fig. 1A. To construct the single-mutant pS1L, the gel-purified 532-nt PCR product from a pDrep1-templated reaction with primers BCV#3(−) and BCV3′end(+) and the gel-purified 397-nt PCR product from a pDrep1-templated reaction with primers 5′S1(−) and Reverse(+) were combined in an overlap extension reaction with primers BCV#3(−) and Reverse(+) to form a product of 786 nt. From this, the 566-nt *Nsi*I-*Mlu*I fragment was used to replace the corresponding region in pDrep1. All other mutant DI RNAs were similarly constructed, except for the use of primers 3′S1(−), 5′S2(−), and 3′S2(−) in the first PCR mutagenesis reaction (resulting in 365-, 377-, and 352-nt

* Corresponding author. Mailing address: Department of Microbiology, University of Tennessee, M409 Walters Life Sciences Building, Knoxville, TN 37996-0845. Phone: (423) 974-4030. Fax: (423) 974-4007. E-mail: dbrian@utk.edu.

† Present address: National Dong-Hwa University, Hualien, Taiwan.

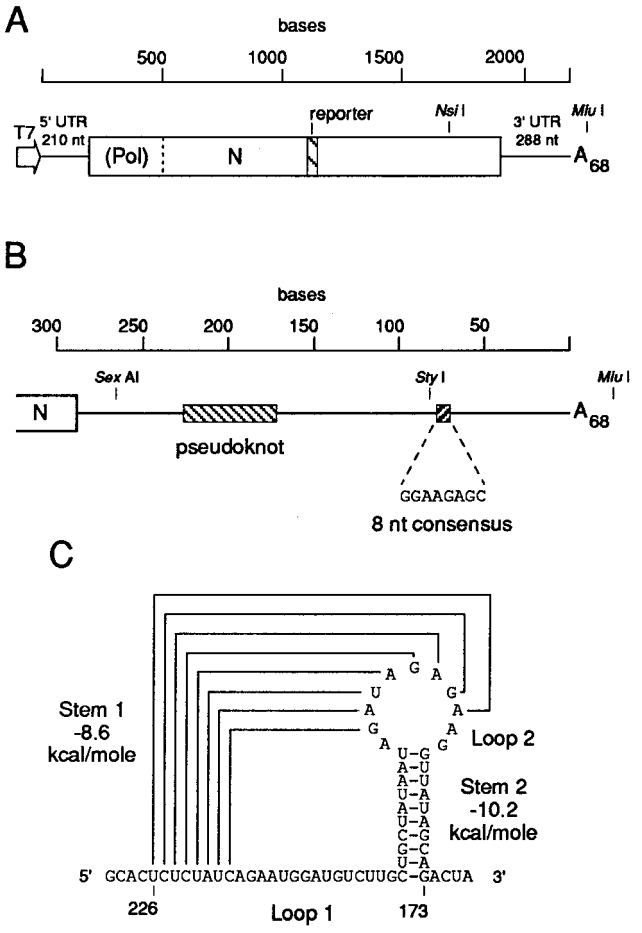


FIG. 1. Proposed pseudoknot and its location within the BCV 3′ UTR. (A) Cloned reporter-containing DI RNA (pDrep1) of BCV used to test the RNA replication function of the pseudoknot. The cloned DI RNA is under control of a T7 RNA polymerase promoter in plasmid vector pGEM3Zf(−) (Promega). The reporter is an in-frame 30-nt sequence from the N gene of TGEV (11). The numbering starts at the 5′ end of the transcript. (B) Schematic location of the pseudoknot and of an 8-nt consensus sequence (32) in the genomic 3′ UTR. The numbering starts at the base of the poly(A) tail. (C) Proposed pseudoknot structure.

products, respectively) to generate the single mutants pS1R, pS2L, and pS2R, respectively, and the use of pS1L and pS2L template DNA in the 532-nt PCR product reaction to generate the double mutants pS1L/R and pS2L/R, respectively. All transferred PCR fragments were completely sequenced to exclude off-site mutations.

Assay for DI RNA replication. The assay for DI RNA replication was performed as previously described (11), except that cells at approximately 3.4×10^6 per 35-mm-diameter dish at 80% confluency were infected with BCV at a multiplicity of 25 PFU per cell and transfected with 600 ng of transcript. For passage of progeny virus, supernatant fluids were harvested at 48 h posttransfection and 500 μl was used to directly infect freshly confluent cells in a 35-mm dish. Following Nonidet P-40 lysis, proteinase K digestion, and ethanol precipitation, a set volume from each of the preparations was centrifuged, and the cytoplasmic RNA (typically 2.5 to 5 μg) was applied to individual lanes in a formaldehyde-agarose gel. Approximately 1 ng of transcript was loaded per lane when used as marker. Northern blots were probed with the oligonucleotide TGEV#8(+) 5′-end labeled with ^{32}P to specific activities ranging from 5.5×10^5 to 8.8×10^5 cpm/pmol (Cerenkov counts) and exposed to Kodak XAR-5 film in the presence of an intensifying screen for 6 to 24 h at −80°C. The replication of pDrep1 transcripts was assayed in parallel in all experiments (not shown).

Total RNA extraction and RT-PCR analysis. Fresh cells were infected with supernatant fluids collected at 48 h posttransfection, and total RNA was isolated with TRIzol reagent (GIBCO BRL) at 6 h postinfection, a time of peak minus-strand RNA synthesis (21). The conditions for cDNA synthesis and PCR were those described by the manufacturer of SuperScript II reverse transcriptase (GIBCO BRL). A 20 μl reverse transcription (RT) reaction was performed with approximately 5 μg of total RNA and the primer BCVDI420(−), complementary to sequences in gene 1a (polymerase) and specific to minus-strand DI and genomic RNAs (11). The cDNA then served as template in a nested-PCR strategy with primers located within the 3′ UTR and the DI reporter sequence. The first PCR, with the outer primers BCVDI420(−) and BCV3′end(+), used 2 μl of cDNA solution and consisted of 40 cycles of DNA amplification (30 s at 94°C, 1 min at 55°C, and 2 min at 72°C), and the second PCR, with the inner primers TGEV#7(−) and BCV3′UTR(+), used 5 μl of the first PCR mixture and consisted of 30 cycles of DNA amplification (30 s at 94°C, 1 min at 55°C, and 1 min at 72°C). Both PCR mixtures contained *Taq* DNA polymerase (Promega) in a reaction volume of 50 μl, and the reactions included an initial denaturation for 3 min at 94°C and a final extension for 10 min at 72°C. A PCR product with the expected size of 938 bp was gel purified and used in a ligation reaction as specified in the instructions supplied with the TOPO XL PCR cloning kit (Invitrogen).

Enzymatic probing of in vitro-transcribed RNA. To synthesize a short RNA encompassing the pseudoknot region, a stretch of the 3′ UTR was placed under the T7 promoter. pDrep3 (11), a DI RNA with a 5′ terminus identical to the first 22 nt from the multiple-cloning site of pGEM3Z (Promega), was cut with *Eco*RI and *Sex*AI (Fig. 1B), filled in with T4 DNA polymerase (New England Biolabs), and religated. The resulting construct, p3NS, was confirmed by sequencing through the junction region. RNA transcribed in vitro from *Sty*I-linearized p3NS with T7 RNA polymerase (Promega) contained 9 nt derived from the vector plus 188 nt of the BCV 3′ UTR.

The method of enzymatic probing was performed as previously described (12), except that 2 μg of template DNA was included in the transcription reaction and RNA was renatured by slow cooling at room temperature. Digestion reactions with RNase T_1 (GIBCO BRL) (results not shown), RNase T_2 (GIBCO BRL) and RNase V_1 (Pharmacia) were done at 4°C for 30 min. The primer BCV3′end(+), complementary to sequences at the 3′ end of the RNA, was 5′-end labeled with ^{32}P, annealed to one-half of the treated RNA, and extended with avian myeloblastosis virus reverse transcriptase (Promega). The products were separated on a 6% polyacrylamide–8.33 M urea gel.

Synthetic oligonucleotides. The oligonucleotides used in this study are described in Table 1.

RESULTS

A predicted pseudoknot in the 3′ UTR of BCV is highly conserved among coronaviruses. To explore the RNA-folding pattern of the 288-nt 3′ UTR of BCV, computer-assisted secondary-structure analysis was performed with the Tinoco algorithm (63). This thermodynamically based approach predicted the existence of two stem regions between 226 and 173 nt from the base of the poly(A) tail (Fig. 1B) with high negative free-energy values ($\Delta G = -8.6$ and -10.2 kcal[1 cal = 4.184 J]/mol for stems 1 and 2, respectively) and the potential to form a hairpin-type (or classical) pseudoknot (Fig. 1C) (68). A classical pseudoknot is a tertiary interaction involving base pairing between a single-stranded region in a hairpin loop and unpaired bases outside of the loop (reviewed in references 47 and 62). When folded, the base-pairing loop region becomes adjacent to the other stem, leading to coaxial stacking of the two stem regions and formation of a quasi-continuous double helix. The proposed pseudoknot is defined by stems 1 and 2 (8 and 10 bp, respectively), connecting loops 1 and 2 (15 and 2 nt, respectively), and a single intervening nucleotide between the two stems.

The pseudoknot identified in this analysis differs from a classical hairpin pseudoknot in two respects. (i) In a classical pseudoknot, loop 2 is generally larger than loop 1. Loop 1 crosses the deep and narrow major groove of stem 2, and loop 2 crosses the shallow and wide minor groove of stem 1. The BCV pseudoknot has a relatively lengthy loop 1, implying that these sequences have a biological relevance. On the other hand, loop 2, 3 nt long with the fraying of the 182G-194U base pair at the top of stem 2, is not sufficient to bridge the minor groove of an 8-bp (stem 1) A-form helix (48). (ii) The stem regions in a classical pseudoknot are contiguous. The BCV pseudoknot has an insertion of 1 nt, A193, between the two stem regions, which may prevent a linear arrangement of the stems. The presence of the intervening nucleotide, along with the steric constraints caused by the short length of loop 2,

TABLE 1. Synthetic oligonucleotides in this study

Oligonucleotide[a]	Sequence[b]	Binding region[c]
BCVDI420(−)	5′GTTGTGTGCAGTCTAGCCTAATAC3′	419–442
TGEV#8(+)	5′CATGGCACCATCCTTGGCAACCCAGA3′	1099–1128
TGEV#7(−)	5′TCTGGGTTGCCAAGGATGGTGCCATG3′	1099–1128
BCV#3(−)	5′TGGGAATCTTGACGAGCCCCAGAAGG3′	1542–1567
5′S1(−)	5′AAGGCACTCTTGTACAGAATGGATG3′	1931–1955
5′S2(−)	5′GGATGTCTTGCTCGATTAATAGATAGAG3′	1951–1978
3′S1(−)	5′GCTATAATAGTACAAGCAGGTTATAGC3′	1963–1989
3′S2(−)	5′GAGAAGGTTAATCGAGACTATAGATTAATTAG3′	1976–2007
BCV3′UTR(+)	5′CCAACACTATACATTACCAC3′	2021–2040
BCV3′end(+)	5′TCGGCAATTACTTCCGCAAG3′	2054–2073
Reverse(+)	5′CACAGGAAACAGCTATGACC3′	

[a] Oligonucleotides bind to either plus-strand RNA, as indicated by (+), or to minus-strand RNA, as indicated by (−).
[b] Underlined bases indicate differences from the genomic sequence.
[c] Numbers correspond to the pDrep1 plus-strand sequence, except for oligonucleotides TGEV#8(+) and TGEV#7(−), for which numbers refer to the TGEV N gene plus-strand sequence (11). Reverse(+) binds to pGEM3Zf(−) DNA just downstream of the multiple cloning region.

probably results in a bent conformation of the pseudoknot with stems 1 and 2 tilting toward each other (58).

To determine whether this pseudoknotted structure is supported by phylogenetic analysis, the 3′ UTRs of all sequenced coronaviruses were examined. These comparisons revealed a similar pseudoknot in the same relative location for each coronavirus. Among the mammalian coronaviruses (Fig. 2A), structure at the secondary and tertiary levels is maintained despite differences in RNA sequences. Only minor variations are found in stem and loop lengths, and a number of residues are absolutely conserved (note the consensus structure in Fig. 2A). An analogous structure in the avian coronavirus infectious bronchitis virus (IBV) (Fig. 2B) diverges more considerably in stem and loop dimensions but retains some of the conserved features of the mammalian coronaviruses (e.g., a consensus motif nAnnnnGGnnnnnnnnnn in IBV loop 2 found in mammalian coronavirus loop 1). Taken together, these observations led to the working hypothesis that the pseudoknot functions in the coronavirus life cycle.

Base pairing in both stems of the pseudoknot is required for BCV DI RNA replication. Previous studies have described the cloning of a 2.2-kb reporter-containing BCV DI RNA replicon (pDrep1) (11). The DI RNA is a fusion between the 5′ and 3′ termini of the virus genome, resulting in a single fused open reading frame (Fig. 1A). When synthetic uncapped transcripts from *Mlu*I-linearized pDrep1 were transfected into BCV (helper virus)-infected cells, replication was observed as monitored by Northern analysis with a strand-specific probe for the reporter sequence (11). Plus-strand DI RNA accumulated in transfected cells with kinetics paralleling that of helper virus genome, and it appeared during passage 1 of progeny virus on fresh cells.

To test whether the pseudoknot plays a role in RNA replication, site-directed mutagenesis was undertaken to disrupt and then restore base pairing in both stems in pDrep1. Multiple base substitutions were designed to create mismatches predicted to destroy the thermodynamic stability of each stem (data not shown). The mutations were derived from the analogous stem in porcine transmissible gastroenteritis virus (TGEV) (for stem 1 single mutants pS1L and pS1R) and from the other side of the stem (for stem 2 single mutants pS2L and pS2R). Double mutants for stems 1 and 2 (pS1L/R and pS2L/R, respectively) incorporating both sets of mutations were also generated, and all mutant constructs were assayed for replication.

The effects of these mutations were demonstrated by Northern analysis and are shown in Fig. 3. The single mutants with disrupted stems showed no evidence of replication. On the other hand, the double mutants with restored stems replicated at or near wild-type levels and became packaged into virions as evidenced by successful passaging. A lower negative free-energy value ($\Delta G = -5.8$ kcal/mol) for the 7-bp TGEV-mimicked stem in pS1L/R and alterations in the highly conserved bases A198, G195, and U178 in the restored stem in pS2L/R were tolerated by the DI RNA. The latter finding indicates that the primary sequence in these regions is not a requirement for replication. In summary, these experiments establish a correlation between the maintenance of the two stem regions in the pseudoknot and BCV DI RNA replication.

The mutated sequences in pS1L/R and pS2L/R are retained after a viral passage. To evaluate the integrity of the mutations in the replicating transcripts from pS1L/R and pS2L/R, RT-PCR analysis was performed on minus-strand templates in total RNA extracted after the first viral passage. This approach should ensure that the product was derived from only molecules that had undergone replication and packaging. A DI RNA-specific product was amplified and cloned, and five insert-containing colonies for each construct were sequenced across the region of the pseudoknot (data not shown). The number of colonies containing matched pairs of mutations for pS1L/R and pS2L/R were two and four, respectively. The remaining sequences were of wild-type origin, indicating that the VP1 bands on the Northern blots (Fig. 3) represent a mixture of DI RNA molecules, presumably the consequence of recombination events between the helper virus and the transfected DI RNA. The results also show that the double mutants undergo replication and become packaged in supernatant virus.

Enzymatic probing yields evidence for both open and closed forms of the pseudoknot. To obtain direct biochemical evidence for the presence of the pseudoknot, the structure of an in vitro-transcribed RNA was probed by enzymatic treatment. The 197-nt RNA substrate, including 40 nt upstream and 94 nt downstream of the pseudoknot, was digested with RNases specific for single-stranded regions (RNase T_1 [G specific] and RNase T_2 [nonspecific with a preference for A]) and helical regions (cobra venom RNase V_1). The positions of cleavage sites in the RNA were determined by analyzing the primer extension products in parallel with a dideoxy sequencing ladder generated from untreated RNA and the same primer. Figure 4A, an autoradiograph of a representative gel, shows that increasing the amount of enzyme in the digestion mixture yields specific banding patterns. A summary of all enzymatic

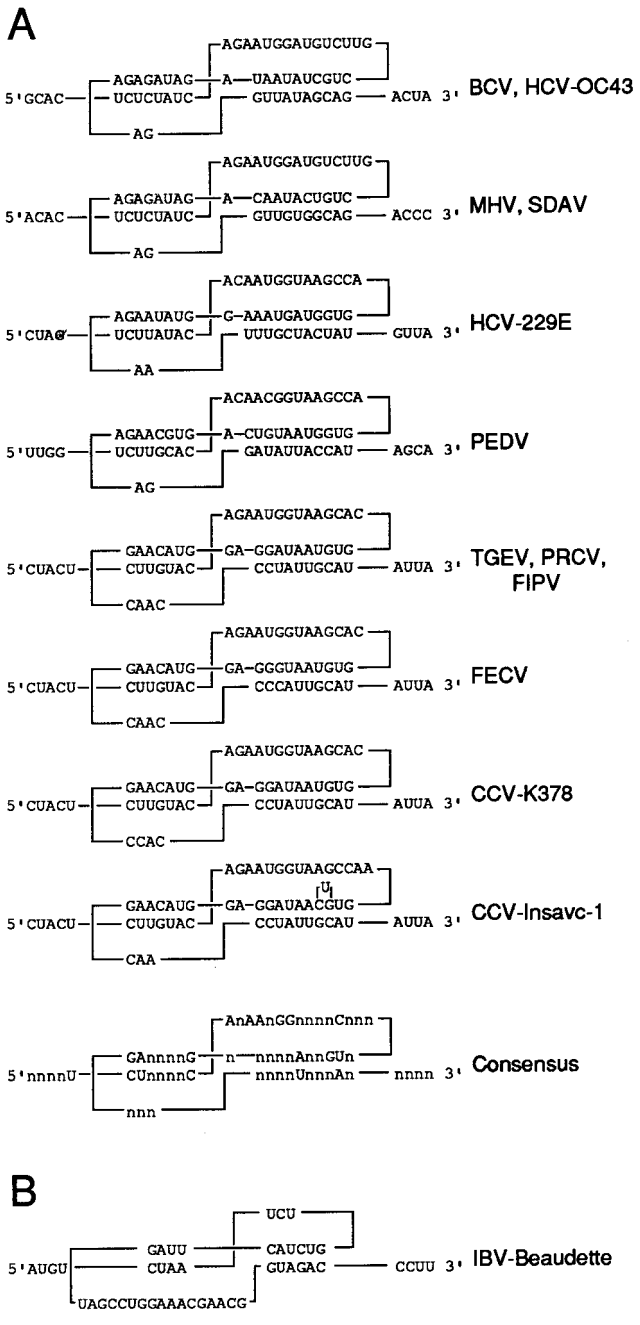


FIG. 2. Comparison of the phylogenetically conserved pseudoknots among coronaviruses. Stems 1 and 2 are shown on the left and right, respectively. (A) Members of coronavirus serogroups 1 (HCV-229E [54], PEDV [9], TGEV [28], PRCV [49], FIPV [13], FECV [67] and CCV [22, 67]), and 2 (BCV [32], HCV-OC43 [27], MHV [44], and SDAV [30]) are shown. (B) IBV (6), the only member of coronavirus serogroup 3, is shown. Abbreviations: HCV, human coronavirus; SDAV, rat sialodacryoadenitis virus; PEDV, porcine epidemic diarrhea virus; PRCV, porcine respiratory coronavirus; FIPV, feline infectious peritonitis virus; FECV, feline enteric coronavirus; CCV, canine coronavirus.

probing data superimposed on the pseudoknot is given in Fig. 4B.

Three major areas were susceptible to RNase V_1: those between C225 and G217, G209 and C207, and C200 and A196. RNase V_1 recognizes RNA sequences in a double-stranded region (a canonical helix) and, on occasion, in a single-

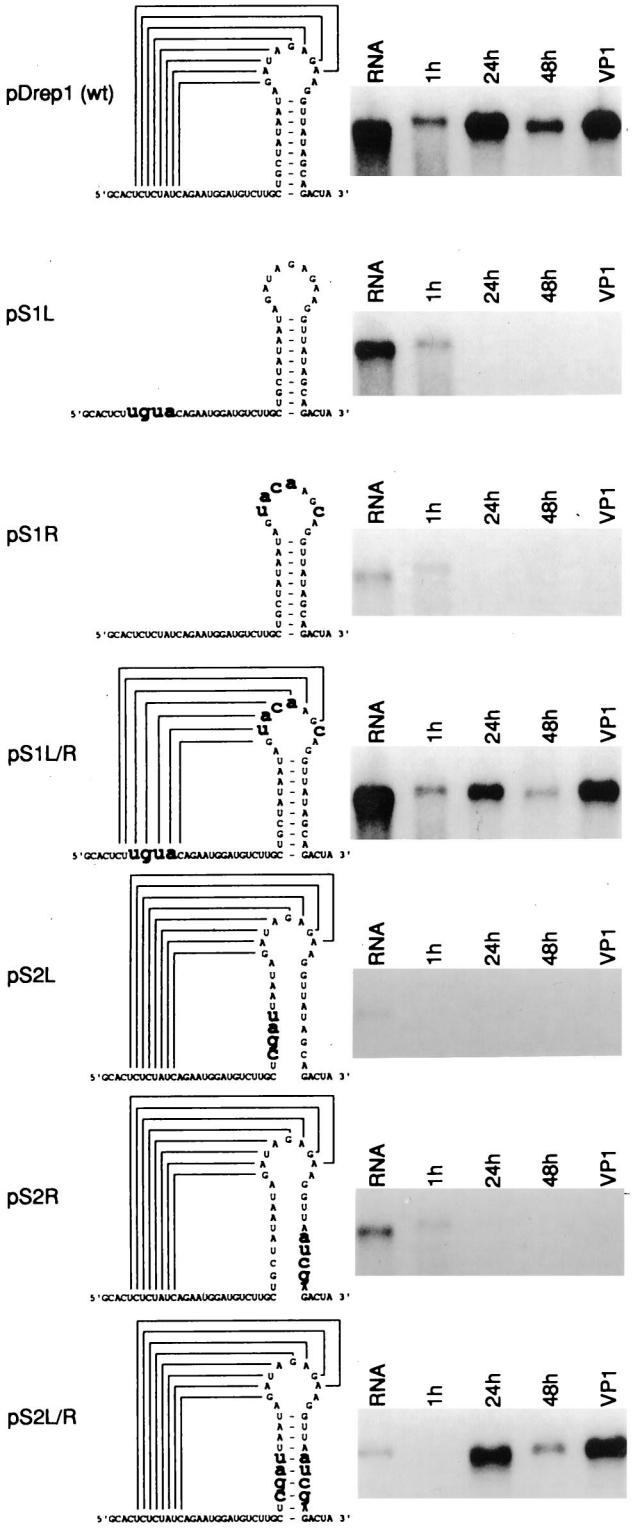
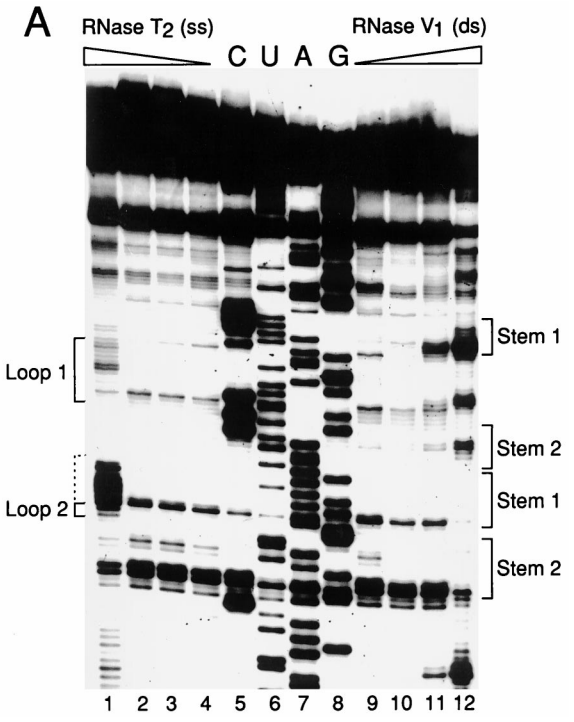


FIG. 3. Importance of stems 1 and 2 in DI RNA replication. Synthetic transcripts of pDrep1 and six mutants were transfected in BCV-infected cells, and cytoplasmic RNA was extracted at 1, 24, and 48 h posttransfection and at 48 h after the first virus passage (VP1). Extracted RNA was separated by electrophoresis on a formaldehyde-agarose gel and probed in a Northern blot with radiolabeled oligonucleotide for detection of the plus strand of the reporter sequence. An increase in the amount of DI RNA after 1 h posttransfection is evidence of replication. RNA shown in the first lane is an aliquot of the in vitro transcript used as marker.



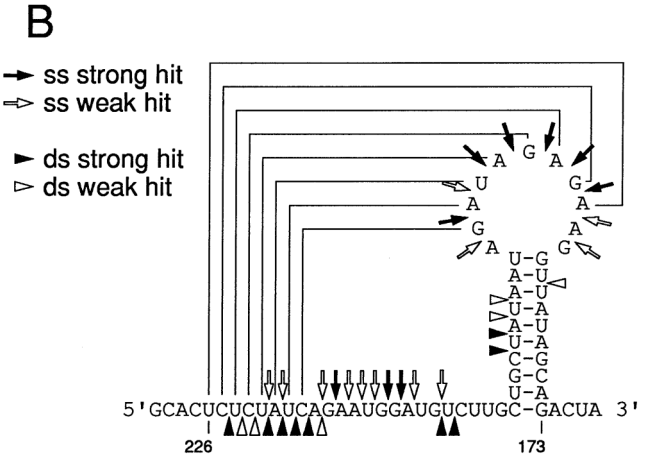


FIG. 4. Enzymatic probing of the pseudoknot in synthetic RNA transcripts. (A) Electrophoretic analysis of digestion products. Lanes: 1 to 4, RNase T_2 digestion with 0.65, 0.065, 0.0065, and 0 U, respectively; 5 to 8, sequencing ladder generated from the same primer; 9 to 12, RNase V_1 digestion with 0, 0.01, 0.1, and 1.0 U, respectively. The region encompassed by the dashed bracket may show both single-stranded and double-stranded properties (see the text). Results from digestion with RNase T_1 are not shown. ss, single-stranded; ds, double-stranded. (B) Schematic summary of enzymatic probing.

stranded region adopting an approximately helical conformation (37). An interpretation that the first and third of these sets of cleavages occur in double-helical RNA is consistent with the base pairing predicted for both stems 1 and 2. The second set might arise during an altered conformation of the pseudoknot (postulated below). The reactive sites between C219 and G217, located outside stem 1, may result from the ability of RNase V_1 to cleave a stacked single-stranded region immediately adjacent to a stem (1). Note that only one strong banding pattern is observed for each stem. An explanation for this result is that RNase V_1 cleavage on the 5′ side of a stem is generally greater than that on the 3′ side of a stem (1).

RNase T_2-sensitive areas, on the other hand, were found between A218 and U210 and between A193 and G183. Digestion with RNase T_1 yielded similar results, with strong cleavages between G213 and A211, between G188 and A187, and between G186 and A185 (data not shown). These single-stranded regions include not only loops 1 and 2 but also the side of stem 1 between G192 and A185, in apparent contradiction to the predicted pseudoknot. These data are most compatible with a model in which the base pairings in stem 1 are present in a "breathing" (open and closed) interaction. Therefore, the pseudoknot may exist in equilibrium with the hairpin formed by stem 2. An "open" pseudoknot also may account for the area of RNase V_1 reactivity between G209 and C207 by allowing these bases to interact with another single-stranded region.

An RNA sequence able to form a pseudoknot also has the potential to form two hairpin structures (stem 1 alone or stem 2 alone). Experiments with oligonucleotides have provided evidence that pseudoknots with coaxially stacked stems are only marginally more stable than either of the constituent hairpins and that equilibria among secondary and tertiary structures are sensitive to changes in ionic conditions, temperature, loop sequence, and loop size (69). A decrease in the size of pseudoknot loop 2, for example, shifts the equilibria among conformations by stabilizing the stem 2 hairpin relative to the pseudoknot. In the proposed pseudoknot, the small size of loop 2 (2 nt) and the higher negative free energy of stem 2 (compared to stem 1) potentially contribute to the observed labile conformation. In addition, the single nucleotide at the stem-stem junction, A193, may not allow coaxial stacking of the stems, resulting in a greater degree of breathability (58).

DISCUSSION

From the experiments presented here, we conclude that a phylogenetically conserved hairpin-type pseudoknot in the 3′ UTR of BCV is a *cis*-acting signal for DI RNA replication and, by extension, viral genome replication. We postulate, furthermore, that the structure functions in the plus strand, since a large number of G · U pairs are found in stem 2 of the mammalian coronaviruses (ranging from one in BCV to three in human coronavirus 229E [Fig. 2]) and bases on the minus strand fail to covary to maintain base pairings. Consistent with our conclusions are studies on DI RNA of mouse hepatitis virus (MHV), another serogroup 2 coronavirus, for which it has been shown that the minimal sequences at the 3′ end of the genome required for replication of DIssE (29), DIssF (34), and MIDI (66) include the corresponding region of the pseudoknot. In addition, deletions that remove regions within the pseudoknot have been found to abolish DI RNA replication (29, 34). These mutant constructs, however, also remove all or part of a proposed signaling bulged stem-loop mapping immediately upstream of the pseudoknot (24) and eliminate a 26-nt sequence downstream of the pseudoknot shown to be important in both RNA-protein binding and DI RNA replication (36, 70). Except for a preliminary report of this work (68), a pseudoknotted structure has not been identified as a coronavirus 3′-UTR RNA replication signal.

Although 3′-UTR pseudoknotted tertiary structures in plus-strand animal RNA viruses have been postulated to function as regulatory elements in RNA replication, to our knowledge the only other example of such a structure supported by mutational analyses is the phylogenetically conserved pseudoknot of

enteroviruses (40, 41, 46). In this case, however, the pseudoknot was noted not to be a conventional one, since relatively long-distance loop-loop interactions are needed to form the stems. How this element functions mechanistically in enterovirus RNA replication was not determined.

How might the BCV 3′ UTR pseudoknot exert an effect on RNA synthesis? Despite its superficial similarity to the well-characterized hairpin-type pseudoknot active in the ribosomal frameshifting at the coronavirus gene 1a/1b junction (60), there is no indication from inspection of the sequence that it is involved in a similar frameshifting event or stop codon readthrough that might yield a novel *cis*-acting protein. It also seems unlikely, based on a recent demonstration that the minimal 3′-terminal sequence requirement for minus-strand synthesis in MHV DI RNA is approximately 50 nt (exclusive of the poly (A) tail) (35) and on the finding that the 3′ UTRs of BCV and MHV DI RNAs can be exchanged without loss of replicability (unpublished data), that the pseudoknot is part of the signal directing minus-strand synthesis.

An answer to the question of 3′-UTR pseudoknot function may lie in precedents established by studies of positive-strand plant RNA viruses. Two separate but poorly understood regulatory pathways have been described. In the first, pseudoknots mapping just downstream of the open reading frame in the nonpolyadenylated tobacco mosaic virus genome (the so-called upstream pseudoknots in tobacco mosaic virus) have been shown to functionally mimic the poly(A) tail as an enhancer of translation (17). Inasmuch as translation is a *cis*-acting requirement for RNA replication in many plus-strand RNA viruses (43), regulation of translation could be a means of controlling replication. The replication of BCV DI RNA is also strictly dependent on the translatability of its open reading frame (10), making pseudoknot-mediated translational regulation an attractive model to test.

In the second example, the 3′-terminal pseudoknots in plant virus genomes comprise the acceptor arm in *cis*-acting tRNA-like structures (TLSs) (reviewed in reference 16). The mechanistic contribution of the TLSs to genome replication is not understood, but one hypothesis envisions an interaction with components of the translation apparatus, such as tRNA synthetases and elongation factors, to form the enzymatically active RNA replication complex (18). The coronavirus 3′ UTR pseudoknot, therefore, may function as an independent upstream element to recruit protein factors as part of a replication complex. It should be noted that a tRNA-like element need not be aminoacylatable or even 3′-terminal for recognition by its cognate synthetase, as shown for the TLS within the upstream region of the threonyl-tRNA synthetase mRNA from *Escherichia coli* (reviewed in reference 39). Also, the identity features of a canonical tRNA molecule or TLS that signal synthetase recognition can be contained entirely within the acceptor arm (16). It will therefore be important to determine whether the coronavirus 3′ UTR pseudoknot can be recognized by cellular synthetases and elongation factors.

Three-dimensional modeling and further mutagenesis studies will be required to delineate more precisely the structure and function of the pseudoknot. In addition, it will be important to determine whether similar tertiary structural elements function in the genomes of toroviruses and arteriviruses, the other genera within the order *Nidovirales*.

ACKNOWLEDGMENTS

We thank Jennifer Black and Savithra Senanayake for valuable discussions and David Draper (Johns Hopkins University) and Cornelis Pleij (University of Leiden) for critical comments on the manuscript.

This work was supported by Public Health Service grant AI14367 from the National Institutes of Health and by funds from the University of Tennessee College of Veterinary Medicine Center of Excellence Program for Livestock Diseases and Human Health.

REFERENCES

1. **Auron, P. E., L. D. Weber, and A. Rich.** 1982. Comparison of transfer ribonucleic acid structures using cobra venom and S_1 endonucleases. Biochemistry **21:**4700–4706.
2. **Baer, M. L., F. Houser, L. S. Loesch-Fries, and L. Gehrke.** 1994. Specific RNA binding by amino-terminal peptides of alfalfa mosaic virus coat protein. EMBO J. **13:**727–735.
3. **Baric, R. S., S. A. Stohlman, and M. M. C. Lai.** 1983. Characterization of replicative intermediate RNA of mouse hepatitis virus: presence of leader RNA sequences on nascent chains. J. Virol. **48:**633–640.
4. **Blackwell, J. L., and M. A. Brinton.** 1995. BHK cell proteins that bind to the 3′ stem-loop structure of the West Nile virus genome RNA. J. Virol. **69:** 5650–5658.
5. **Blight, K. J., and C. M. Rice.** 1997. Secondary-structure determination of the conserved 98-base sequence at the 3′ terminus of hepatitis C virus genome RNA. J. Virol. **71:**7345–7352.
6. **Boursnell, M. E. G., M. M. Binns, I. J. Foulds, and T. D. K. Brown.** 1985. Sequences of the nucleocapsid genes from two strains of avian infectious bronchitis virus. J. Gen. Virol. **66:**573–580.
7. **Brian, D. A., D. E. Dennis, and J. S. Guy.** 1980. Genome of porcine transmissible gastroenteritis virus. J. Virol. **34:**410–415.
8. **Brian, D. A., and W. J. M. Spaan.** 1997. Recombination and coronavirus defective interfering RNAs. Semin. Virol. **8:**101–111.
9. **Bridgen, A., M. Duarte, K. Tobler, H. Laude, and M. Ackermann.** 1993. Sequence determination of the nucleocapsid protein gene of the porcine epidemic diarrhoea virus confirms that this virus is a coronavirus related to human coronavirus 229E and porcine transmissible gastroenteritis virus. J. Gen. Virol. **74:**1795–1804.
10. **Chang, R.-Y., and D. A. Brian.** 1996. *cis*-requirement for N-specific protein sequence in bovine coronavirus defective interfering RNA replication. J. Virol. **70:**2201–2207.
11. **Chang, R.-Y., M. A. Hofmann, P. B. Sethna, and D. A. Brian.** 1994. A *cis*-acting function for the coronavirus leader in defective interfering RNA replication. J. Virol. **68:**8223–8231.
12. **Chang, R.-Y., R. Krishnan, and D. A. Brian.** 1996. The UCUAAAC promoter motif is not required for high-frequency leader recombination in bovine coronavirus defective interfering RNA. J. Virol. **70:**2720–2729.
13. **DeGroot, R. J., A. C. Andeweg, M. C. Horzinek, and W. J. Spaan.** 1988. Sequence analysis of the 3′-end of the feline coronavirus FIPV 79-1146 genome: comparison with the genome of porcine coronavirus TGEV reveals large insertions. Virology **167:**370–376.
14. **Felden, B., C. Florentz, R. Giege, and E. Westhof.** 1996. A central pseudoknotted three-way junction imposes tRNA-like mimicry and the orientation of three 5′ upstream pseudoknots in the 3′ terminus of tobacco mosaic virus RNA. RNA **2:**201–212.
15. **Felden, B., C. Florentz, E. Westhof, and R. Geige.** 1993. Non-canonical substrates of aminoacyl-tRNA synthetases: the tRNA-like structure of brome mosaic virus genomic RNA. Biochimie **75:**1143–1157.
16. **Florentz, C., and R. Giege.** 1995. tRNA-like structures in plant viral RNAs, p. 141–164. *In* D. Soll and U. L. RajBhandary (ed.), tRNA: structure, biosynthesis, and function. American Society for Microbiology, Washington, D.C.
17. **Gallie, D. R., and V. Walbot.** 1990. RNA pseudoknot domain of tobacco mosaic virus can functionally substitute for a poly(A) tail in plant and animal cells. Genes Dev. **4:**1149–1157.
18. **Giege, R.** 1996. Interplay of tRNA-like structures from plant viral RNAs with partners of the translation and replication machineries. Proc. Natl. Acad. Sci. USA **93:**12078–12081.
19. **Giege, R., C. Florentz, and T. W. Dreher.** 1993. The TYMV tRNA-like structure. Biochimie **75:**569–582.
20. **Hofmann, M. A., and D. A. Brian.** 1991. The 5-prime end of coronavirus minus-strand RNAs contains a short poly(U) tract. J. Virol. **65:**6331–6333.
21. **Hofmann, M. A., P. B. Sethna, and D. A. Brian.** 1990. Bovine coronavirus mRNA replication continues throughout persistent infection in cell culture. J. Virol. **64:**4108–4114.
22. **Horsburgh, B. C., I. Brierley, and T. D. Brown.** 1992. Analysis of a 9.6 kb sequence from the 3′ end of canine coronavirus genomic RNA. J. Gen. Virol. **73:**2849–2862.
23. **Horton, R. M., Z. Cai, S. N. Ho, and L. R. Pease.** 1990. Gene splicing by overlap extension: tailor made genes using the polymerase chain reaction. BioTechniques **8:**528–535.
24. **Hsue, B., and P. S. Masters.** 1997. A bulged stem-loop structure in the 3′ untranslated region of the genome of the coronavirus mouse hepatitis virus is essential for replication. J. Virol. **71:**7567–7578.
25. **Ito, T., and M. M. C. Lai.** 1997. Determination of the secondary structure of and cellular protein binding to the 3′-untranslated region of the hepatitis C virus RNA genome. J. Virol. **71:**8698–8706.

26. **Jacobson, S. J., D. A. M. Konings, and P. Sarnow.** 1993. Biochemical and genetic evidence for a pseudoknot structure at the 3′ terminus of the poliovirus RNA genome and its role in viral RNA amplification. J. Virol. **67:**2961–2971.
27. **Kamahora, T., L. H. Soe, and M. M. C. Lai.** 1989. Sequence analysis of nucleocapsid gene and leader RNA of human coronavirus OC43. Virus Res. **12:**1–9.
28. **Kapke, P. A., and D. A. Brian.** 1986. Sequence analysis of the porcine transmissible gastroenteritis coronavirus nucleocapsid protein gene. Virology **151:**41–49.
29. **Kim, Y.-N., Y. S. Jeong, and S. Makino.** 1993. Analysis of *cis*-acting sequences essential for coronavirus defective interfering RNA replication. Virology **197:**53–63.
30. **Kunita, S., M. Mori, and E. Terada.** 1993. Sequence analysis of the nucleocapsid protein gene of rat coronavirus SDAV-681. Virology **193:**520–523.
31. **Lai, M. M. C., and D. Cavanagh.** 1997. The molecular biology of coronaviruses. Adv. Virus Res. **48:**1–100.
32. **Lapps, W., B. G. Hogue, and D. A. Brian.** 1987. Sequence analysis of the bovine coronavirus nucleocapsid and matrix protein genes. Virology **157:**47–57.
33. **Leathers, V., R. Tanguay, M. Kobayashi, and D. R. Gallie.** 1993. A phylogenetically conserved sequence within viral 3′ untranslated RNA pseudoknots regulates translation. Mol. Cell. Biol. **13:**5331–5347.
34. **Lin, Y.-J., and M. M. C. Lai.** 1993. Deletion mapping of a mouse hepatitis virus defective interfering RNA reveals the requirement of an internal and discontiguous sequence for replication. J. Virol. **67:**6110–6118.
35. **Lin, Y.-J., C.-L. Liao, and M. M. C. Lai.** 1994. Identification of the *cis*-acting signal for minus-strand RNA synthesis of a murine coronavirus: implications for the role of minus-strand RNA in RNA replication and transcription. J. Virol. **68:**8131–8140.
36. **Liu, Q., W. Yu, and J. L. Leibowitz.** 1997. A specific host cellular protein binding element near the 3′ end of mouse hepatitis virus genomic RNA. Virology **232:**74–85.
37. **Lowman, H. B., and D. E. Draper.** 1986. On the recognition of helical RNA by cobra venom V_1 nuclease. J. Biol. Chem. **261:**5396–5403.
38. **Luytjes, W.** 1995. Coronavirus gene expression: Genome organization and protein synthesis, p. 33–54. *In* S. G. Siddell (ed.), The Coronaviridae. Plenum Publishing Corp., New York, N.Y.
39. **Mans, R. M. W., C. W. A. Pleij, and L. Bosch.** 1991. tRNA-like structures: structure, function and evolutionary significance. Eur. J. Biochem. **201:**303–324.
40. **Melchers, W. J. G., J. G. J. Hoenderop, H. J. Bruins Slot, C. W. A. Pleij, E. V. Pilipenko, V. I. Agol, and J. M. D. Galama.** 1997. Kissing of the two predominant hairpin loops in the coxsackie B virus 3′ untranslated region is the essential structural feature of the origin of replication required for negative-strand RNA synthesis. J. Virol. **71:**686–696.
41. **Mirmomeni, M. H., P. J. Hughes, and G. Stanway.** 1997. An RNA tertiary structure in the 3′ untranslated region of enteroviruses is necessary for efficient replication. J. Virol. **71:**2363–2370.
42. **Nakhasi, H. L., T. A. Rouault, D. J. Haile, T.-Y. Liu, and R. D. Klausner.** 1990. Specific high-affinity binding of host cell proteins to the 3′ region of rubella virus RNA. New Biol. **2:**255–264.
43. **Novak, J. E., and K. Kirkegaard.** 1994. Coupling between genome translation and replication in an RNA virus. Genes Dev. **8:**1726–1737.
44. **Parker, M. M., and P. S. Masters.** 1990. Sequence comparison of the N genes of five strains of the coronavirus mouse hepatitis virus suggests a three domain structure for the nucleocapsid protein. Virology **179:**463–468.
45. **Pilipenko, E. V., S. V. Maslova, A. N. Sinyakov, and V. I. Agol.** 1992. Towards identification of *cis*-acting elements involved in the replication of enterovirus and rhinovirus RNAs: a proposal for the existence of tRNA-like terminal structures. Nucleic Acids Res. **20:**1739–1745.
46. **Pilipenko, E. V., K. Poperechny, S. V. Maslova, W. J. G. Melchers, H. J. Bruins Slot, and V. I. Agol.** 1996. *cis*-element, *oriR*, involved in the initiation of (−) strand poliovirus RNA: a quasi-globular multi-domain RNA structure maintained by tertiary ('kissing') interactions. EMBO J. **15:**5428–5436.
47. **Pleij, C. W. A.** 1994. RNA pseudoknots. Curr. Opin. Struct. Biol. **4:**337–344.
48. **Pleij, C. W. A., K. Rietveld, and L. Bosch.** 1985. A new principle of RNA folding based on pseudoknotting. Nucleic Acids Res. **13:**1717–1731.
49. **Rasschaert, D., M. Duarte, and H. Laude.** 1990. Porcine respiratory coronavirus differs from transmissible gastroenteritis virus by a few genomic deletions. J. Gen. Virol. **71:**2599–2607.
50. **Sawicki, S. G., and D. L. Sawicki.** 1990. Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. J. Virol. **64:**1050–1056.
51. **Sawicki, S. G., and D. L. Sawicki.** 1995. Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. Adv. Exp. Med. Biol. **380:**499–506.
52. **Schaad, M. C., and R. S. Baric.** 1994. Genetics of mouse hepatitis virus transcription: evidence that subgenomic negative strands are functional templates. J. Virol. **68:**8169–8179.
53. **Schochetman, G., R. H. Stevens, and R. W. Simpson.** 1977. Presence of infectious polyadenylated RNA in the coronavirus avian bronchitis virus. Virology **77:**772–782.
54. **Schreiber, S. S., T. Kamahora, and M. M. C. Lai.** 1989. Sequence analysis of the nucleocapsid protein gene of human coronavirus 229E. Virology **169:**142–151.
55. **Senanayake, S. D., and D. A. Brian.** 1995. Precise large deletions by the PCR-based overlap extension method. Mol. Biotechnol. **4:**13–15.
56. **Sethna, P. B., M. A. Hofmann, and D. A. Brian.** 1991. Minus-strand copies of replicating coronavirus mRNAs contain antileaders. J. Virol. **65:**320–325.
57. **Sethna, P. B., S.-L. Hung, and D. A. Brian.** 1989. Coronavirus subgenomic minus-strand RNA and the potential for mRNA replicons. Proc. Natl. Acad. Sci. USA **86:**5626–5630.
58. **Shen, L. X., and I. Tinoco.** 1995. The structure of an RNA pseudoknot that causes efficient frameshifting in mouse mammary tumor virus. J. Mol. Biol. **247:**963–978.
59. **Shi, P.-Y., M. A. Brinton, J. M. Veal, Y. Y. Zhong, and W. D. Wilson.** 1996. Evidence for the existence of a pseudoknot structure at the 3′ terminus of the flavivirus genomic RNA. Biochemistry **35:**4222–4230.
60. **Somogyi, P., A. J. Jenner, I. Brierley, and S. C. Inglis.** 1993. Ribosomal pausing during translation of an RNA pseudoknot. Mol. Cell. Biol. **13:**6931–6940.
61. **Spaan, W., H. Delius, M. Skinner, J. Armstrong, P. Rottier, S. Smeekens, B. A. van der Ziejst, and S. G. Siddell.** 1983. Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. EMBO J. **2:**1839–1844.
62. **ten Dam, E., K. Pleij, and D. Draper.** 1992. Structural and functional aspects of RNA pseudoknots. Biochemistry **31:**11665–11676.
63. **Tinoco, I., P. N. Borer, B. Dengler, M. D. Levine, O. C. Uhlenbeck, D. M. Crothers, and J. Gralla.** 1973. Improved estimation of secondary structure in ribonucleic acids. Nat. New Biol. **246:**40–41.
64. **Todd, S., and B. L. Semler.** 1996. Structure-infectivity analysis of the human rhinovirus genomic RNA 3′ non-coding region. Nucleic Acids Res. **24:**2133–2142.
65. **Tsai, C.-H., C.-P. Cheng, C.-W. Peng, B.-Y. Lin, N.-S. Lin, and Y.-H. Hsu.** 1999. Sufficient length of a poly(A) tail for the formation of a potential pseudoknot is required for efficient replication of bamboo mosaic potexvirus RNA. J. Virol. **73:**2703–2709.
66. **van der Most, R. G., W. Luytjes, S. Rutjes, and W. J. M. Spaan.** 1995. Translation but not the encoded sequence is essential for the efficient propagation of defective interfering RNAs of the coronavirus mouse hepatitis virus. J. Virol. **69:**3744–3751.
67. **Vennema, H., J. W. Rossen, J. Wesseling, M. C. Horzinek, and P. J. Rottier.** 1992. Genomic organization and expression of the 3′ end of the canine and feline enteric coronaviruses. Virology **191:**134–140.
68. **Williams, G. D., R.-Y. Chang, and D. A. Brian.** 1995. Evidence for a pseudoknot in the 3′ untranslated region of the bovine coronavirus genome. Adv. Exp. Med. Biol. **380:**511–514.
69. **Wyatt, J. R., J. D. Puglisi, and I. Tinoco.** 1990. RNA pseudoknots: stability and loop size requirements. J. Mol. Biol. **214:**455–470.
70. **Yu, W., and J. L. Leibowitz.** 1995. Specific binding of host cellular proteins to multiple sites within the 3′ end of mouse hepatitis virus genomic RNA. J. Virol. **69:**2016–2023.